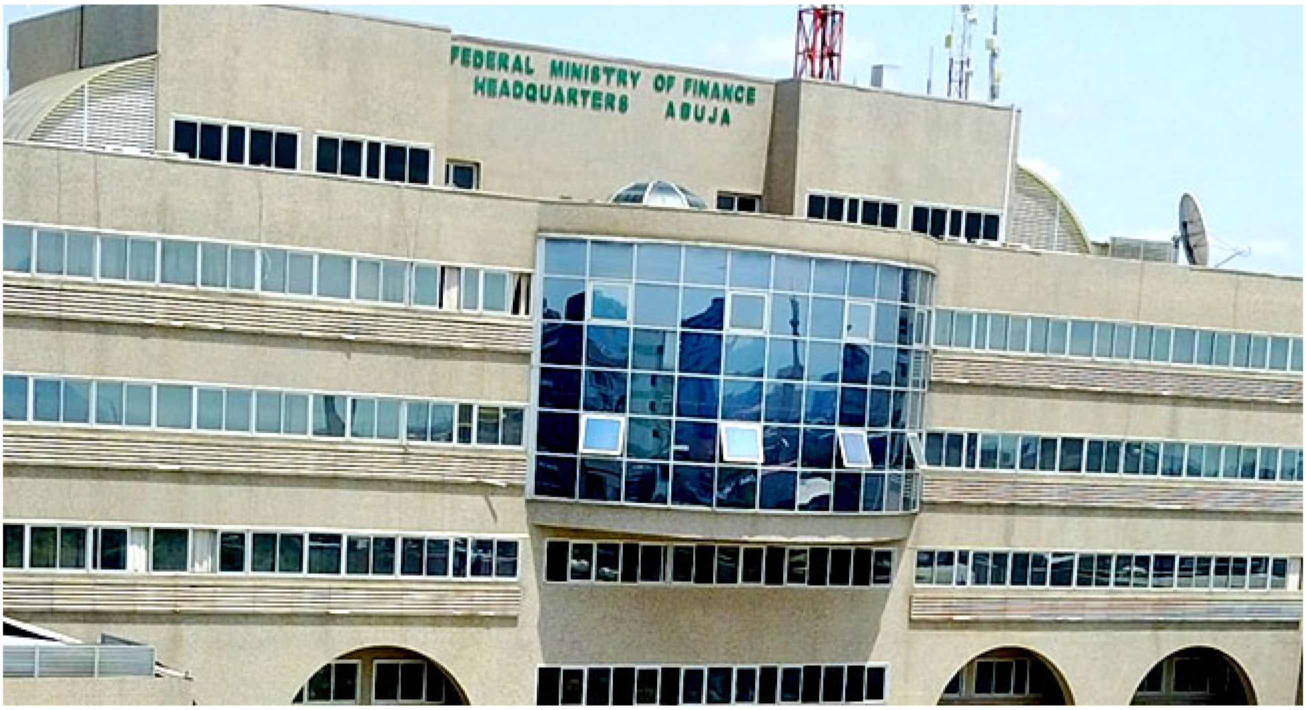
Getting the PFM basics right (A study of PEFA scores awarded over the 2016 and 2011 Frameworks)



By David Fellows and John Leonardo

Introduction

The Public Expenditure and Financial Accountability (PEFA) programme provides a framework for assessing and reporting the strengths and weaknesses of public financial management (PFM). The current 2016 Framework refines the previous 2011 Framework and is structured under a hierarchy of 6 Pillars, 31 Indicators (PIs) and 94 Dimensions. The PEFA Field Guide explains the components of the 2016 Framework and describes how an assessment team should score each dimension on a scale of A to D, a D score representing the lowest level of performance.

An initial assessment of the latest PEFA reports for countries published under the 2016 Framework suggested that many countries were not getting the PFM basics right. This led to a comparison of recent results with those from earlier PEFA reports prepared under the 2011 Framework to examine performance over time and the lessons for PFM improvement that such a comparison may offer (termed the ‘dual study’). It was decided to focus on dimension scores since the demands of PFM can change markedly depending on the aspects of the subject matter under consideration and the evident variations of score for the same country at dimension level within a range of PIs.

It was decided to confine this initial study to the analysis of D scores at the dimension level given the frequency of D scores, the very poor performance they represent and the importance of raising performance to a higher level. The Field Guide requires a D score when: ‘the feature being measured is present at less than the basic level of performance or is absent altogether, or that there is insufficient information to score the dimension’.

For the purpose of this study, D scores include dimensions marked D*, NR and some NA scores where evidence suggests a breakdown in PFM activity. It seemed evident that these attributions are often applied inconsistently and serve to obscure the extent of the poor performance of some countries by avoiding the use of justifiable D scores. A summary of all scores for the 2016 Framework and the dual study evaluations, as discussed in this report, can be accessed at Annex 1.

2016 Framework analysis

The 2016 Framework analysis consisted of the latest published evaluations for the 63 countries for which there were published reports at the time of this study. The D scores represent 32% of all dimension scores in this data set, 39% amongst low-income countries.

D scores were widely distributed throughout the framework with 45 of the 94 dimensions having an above average number of D scores.

The study also defined and assessed the key factors (termed descriptors) that contributed to PFM performance. The results, summarised at Annex 2, suggested that most D scores can be explained by the absence of 'Management Effectiveness', 'Integrity' and in one case of 'High Level Technical Knowledge' although poor "System Design" was another potentially important contributing factor.

Annex 3 provides a full list of the 2016 Framework dimensions and D score data together with the descriptors contributing to each dimension.

Dual framework

Following the results of the 2016 Framework D score study it was decided to undertake a review of 45 countries that have undertaken at least one PEFA evaluation under both the 2011 and 2016 frameworks (the earliest and the latest studies we used for countries with more than two studies). This enabled a country's performance to be compared over a five-year period.

The 2011 and 2016 PEFA frameworks differ in many respects. An equivalence table published by PEFA suggests that the two frameworks can be aligned to 37 "equivalent" dimensions on the basis that the respective dimensions were either "directly comparable" or "indirectly comparable".

The PEFA equivalence table identifies 28 dimensions (or in some cases subsets) from the 2011 framework as "non-comparable (subject only)" to 2016 counterparts suggesting that the dimension descriptions and scoring routines differ markedly while the general area of relevance to the dimensions are similar. This leaves only 37 pairs of comparable dimensions.

On examination, the study team decided that 26 of the 28 pairs of dimensions judged "non-comparable (subject only)" were in fact very similar to the 2016 counterparts, the main difference being the way in which the later guidance is translated into clear-cut scoring criteria but that a good PEFA evaluator should have made reasonably similar judgements for both frameworks when reviewing all but two of these dimensions.

This exercise, therefore, recognises 63 equivalent dimensions while also providing results for PEFA's 37 equivalent dimensions. It is suggested that the D score characteristics of both data sets are sufficiently similar to provide a reasonable validation for the larger 63 dimension equivalence thereby extending the usefulness of inter-framework comparisons. Details of the PEFA and PFMConnect equivalence tables are set out at Annex 4. The dual study of 2016 and 2011

Framework with D score data at dimension level is set out at Annex 5.

The dual study is highly concerning in terms of the lack of improvement amongst those dimensions receiving D scores. These data are further summarised and commented on below.

Table 1: Dual Framework Study Comparison of Results at Dimension Level for the relevant 45 countries						
Data Set	Average D-score across all dimensions		Dimensions with above Average D scores		Dimensions with fewer D scores in 2016 compared with 2011	Countries with fewer D scores in 2016 than 2011
	2016 Framework % (ii)	2011 Framework % (iii)	2016 % (iv)	2011 % (v)	No. (vi)	No. (vii)
(i)						
PEFA's 37 comparable dimensions	26	23	49	46	13	17
PFMConnect's 63 dimensions	28	23	43	41	16	10

The dual framework study reveals a deteriorating performance with most dimensions exhibiting a greater number of D scores in the later evaluations. Only 13 (35%) of dimensions from the 37 dimensions study and 16 (25%) from the 63 dimensions study experienced reductions in D scores between evaluations.

When the dual evaluations for the same country were compared, see Annex 6, it was noted that most countries recorded a higher proportion of D scores for the same dimension in both evaluations demonstrating a reasonably consistent poor performance. A few countries displayed less consistent results.

Few countries in the 63 dimensions set recorded reductions in the number of D scores in 2016 framework results compared with the 2011 framework results. The top performers where significant PFM reform activities had been undertaken between the dual framework studies included: Philippines, Maldives, Mongolia and Tajikistan.

Table 2: Comparing above average D scores for both frameworks

Data Set (i)	Dimensions with above average D scores			
	Total for 2016 framework No. (ii)	Total for 2011 framework No. (iii)	Common to both frameworks) No. (iv)	(%) (v)
PEFA's 37 comparable dimensions	18	18	13	(72)
PFMConnect's 63 dimensions	25	25	20	(77)

NB: Column (iv) shows that of PEFA's 37 equivalent dimensions 13 (35%) have above average D scores common to both frameworks or 20 (32%) for PFMConnect's 63 equivalent dimensions.

The results for the proportion of dimensions with above-average D scores that are common to both framework dimensions sets is concerning. Approximately one third of all dimensions had above-average D scores that were common to both frameworks for the same country for both datasets. In addition, over 70% of the above-average dimensions in both datasets were common to both frameworks showing limited improvement in the worst scoring areas over a five-year period.

Dimensions with regular poor performance are widely distributed (titles in red at Annex 6). This suggests pockets of poor management that remain in place without effective challenge and this is consistent with the descriptor analysis.

Conclusions

This study offers a range of findings that pose questions about the approach, effectiveness and sustainability of PFM reforms instituted by national and subnational governments often in collaboration with development agencies. The concerns about management effectiveness and integrity highlighted in this study must be seen to question the most basic aspects of any organisation.

The study focusses on D score analysis, but it could be useful to extend the analysis to C-level scores where the performance of countries still remains below good international standards. This could reveal new characteristics of national PFM performance and extend the range of analytical techniques applied to performance data.

The data analysis evidences the credibility of PFMConnect's extended 63 dimension equivalence model that offers significant potential for more detailed studies of specific countries or regions.

Further work on descriptors to reveal contributory factors to variations in performance seems worthy of further development.

The failure of some governments to publish PEFA studies in full reinforces concerns about the need for greater attention to integrity. Another improvement that could be readily and widely implemented is legislative scrutiny of audit reports (PI 31).

Recommendations

We recommend that country-specific studies should be undertaken based on PEFA assessment reports (both 2016 Framework studies for the full 94 dimensions and dual studies where the data are available) examining D scores at dimension level to establish potential causes of poor performance and identify ways in which performance may be improved. Issues to consider with respect to areas of poor performance, include:

- The commitment to personnel development and support, including: in-service training, management development, oversight, feedback on performance, and system design.
- The adequacy of transparency and accountability and evidence of corrupt activity.
- The quality of relevant communication and support levels among different departments and units of the finance ministry.
- The reasons for persistently poor or erratic performance and the fit with other findings.
- The observations of managers and staff on reasons for poor performance and barriers to improvement.

We recommend that country studies should be designed as the initial phase of PFM development programmes. In this context, a report by the Swedish International Development Cooperation Agency (SIDA) offers some observations about the conditions for effective PFM reform. These include the importance of change agendas being aligned with Government priorities and the need to treat PFM reform as a learning process with strong emphasis on coordination and systematic evaluation of the activities performed by teams responsible for delivery.

Groups of countries or subnational bodies may wish to collaborate in reform programmes enabling challenges and learning to be shared and systems of mutual support developed. We have previously advocated the use of digital communication as a cost-effective and time-saving way of sharing knowledge and ideas between nations (incl. expert advisors).

Any country, region or development institution wishing to participate in further work in this field is invited to discuss their interest with the authors.

An article based on this study has been published by the IMF's PFM Blog.

PFMConnect is a public financial management consultancy with a particular interest in the use of digital communication to support learning and sharing expertise amongst the international development community.

David Fellows began his career in UK local government where he became President of the Society of Municipal Treasurers and a pioneer of digital government. He has held appointments in the UK Cabinet Office and the National Treasury of South Africa (david.fellows@pfmconnect.com).

John Leonardo is a PFM expert with extensive worldwide experience. He has undertaken PFM assignments in Africa, Asia, the Caribbean and the Pacific where he undertook PEFA assessments. Both authors are directors of PFMConnect, a public financial management consultancy (john.leonardo@pfmconnect.com).

The need to improve the PEFA methodology

We have prepared a SlideShare presentation that recommends improvements to the current Public Expenditure and Financial Accountability (PEFA) methodology for assessing public financial management. You can view the presentation at: http://www.slideshare.net/johnleo/the-need-for-improvement-in-public-expenditure-and-financial-accountability-pefa-assessment-methodology. This presentation supplements the material presented in our recent blog at: Proposals for PEFA reform.

Proposals for PEFA reform

Posted by David Fellows and John Leonardo[1]

The problem

Failings in public financial management span the breadth of the Public Expenditure Financial Accountability (PEFA) scores. Our work suggests that numerous African governments that had very low scores from initial PEFA assessments[2] conducted up to nine years ago for some Performance Indicators (PIs) still present low or failing scores – 'C', 'D+' or even 'D' – for many of the same PIs in the most recent assessments.

PEFA assessments and PEFA-based reforms do not seem to be working for all PIs. Why? There are two main problems.

First, the current PEFA methodology results in assessments often giving little attention to some of the broader institutional causes of poor performance, including:

- **Staff capability:** The selection, availability, training, ambition and management of staff are typically ignored. Even when they are considered, boundaries are often too narrowly drawn around central finance functions.
- **Finance professionalism:** PEFA assessments do not consider the capacity to adapt PFM practice to local characteristics and pressures, to share knowledge and promote essential values.
- **Management:** PEFA assessments rarely question the fitness of the management chain to carry the burden of the finance function and its reform or the commitment of top management and ministers to facilitating this task.
- **ICT capacity:** This is increasingly important – affecting both the ability to operate current processes and an organisation's improvement potential – but completely overlooked.

- **Extraneous policy effects:** requires discipline to avoid concurrent policy changes that will pose unsustainable demands on top of the combined burden of PFM reform activities and the daily routine.
- **The behaviour of politicians and top officials:** In what is the most glaring limitation, corruption at a senior level, capricious decision-taking, unreasonable favouritism and lack of apparent consideration for staff or citizens can have disastrous implications, but features nowhere in the PEFA.

Such omissions can result in PFM reform programs failing to address fundamental problems.

Second, there is a disconnect between analysis and reform, with the two stages often developed by different groups, and reforms often having little relation to key underlying problems.

The very mixed history of PFM reform is a testament to these two problems.

An unacceptable state of affairs

Poor PFM performance is not unique to an African environment or developing nations in general or indeed to developed nations throughout the world. The problem for developing nations is that they are by their very definition less richly resourced. The administrative basis from which they play out their current experience is lower and the safetynets are less robust. They can less readily afford their mistakes and recover from them less rapidly.

The questions that we pose from our analysis are these: (i) is it reasonable to continue to do such a limited evaluation of the overall PFM environment given the evident complexity of the PFM context as we have set out in very brief terms in this note; (ii) given that the PEFA methodology has been in operation for almost ten years and numerous major countries are making relatively little progress with some key PFM functions, is the PEFA methodology and its underlying scope entirely satisfactory; (iii) are not assessments being underutilised in that their scope and the circulation of completed work is often restricted to finance ministries – even where activities outside the central finance function are subject to review there is often limited involvement of staff from such areas in assessment planning, task team membership and post hoc discussion; and (iv) is the best use being made of the insights gained by review teams given that there is no assured linkage between the assessment study and any consequent reform proposals?

A way forward

We believe that at least for those governments in serious difficulty the scope of the PEFA methodology is too narrow and that there must be a more wide-ranging diagnostic review at the PEFA assessment stage that helps concentrate minds on the root causes of serious PFM shortcomings.

It is not uncommon for governments to express doubt about the failings identified in PEFA assessments and reaffirm the validity of plans already made for the future that do not address fundamental problems. Such conclusions are more easily reached when causation is not addressed.

The current PEFA methodology requires the preparation of concept notes to inform decisions concerning the scope of proposed PEFA assessments. We suggest that the merits of undertaking a broader institutional assessment should represent an additional topic to be addressed in the preparation of future concept notes. If evidence of widespread poor performance emerges unexpectedly during an assessment, then the possibility of undertaking a wider institutional review should be considered at that stage.

Whilst it is accepted that political economy factors will have a material influence on PFM outcomes in many countries, that is no reason to ignore them.

Proposals

Once the PEFA diagnosis is complete, and the underlying performance factors laid bare, the same reviewers should be asked to present views on reform priorities, time scales and the reasons for previous reform failures. This work differs from the main assessment process and should therefore be included in an accompanying memorandum.

An extended review process – under which the traditional assessment and an examination of underlying factors are both examined, and reform priorities discussed – would make much better use of the assessment experience as a whole. Such an approach would be more conducive to strong, informed advocacy for prompt and effective decision-taking.

We also believe that assessment reports should be more widely circulated and that a wider group of staff involved in key PFM work activities should be consulted during the assessment process.

Conclusion

The proposed broader-based and extended PEFA assessment methodology falls far short of creating a design tool for PFM reform. It does, however, make better use of the effort and expertise employed and the collaborative environment established during the assessment process. Consequent reform strategies should, therefore, become more effective.

Notes:

[1] The authors are Principals of PFMConnect. They have been engaged on projects in Africa, Asia and the Pacific funded by the World Bank, EU, ADB and DfID, including PEFA assessments. They are grateful to Rajiv Sondhi, Head Loans and Grants at the International Fund for Agricultural Development, who reviewed the draft and offered helpful guidance; nevertheless views expressed remain the responsibility of the authors.

[2] See: www.pefa.org

[3] A slightly abbreviated version of this blog is available at the Devpolicy Blog from the Development Policy Centre based at the Australian National University’s Crawford School of Public Policy: Proposals for PEFA reform